

# Delay Constrained Linear Transmission of Random State Measurements

Onur Tan\*, Deniz Gündüz†, Jesús Gómez Vilardebó\*

\*Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Barcelona, Spain.

†Department of Electrical and Electronic Engineering, Imperial College London, London, UK.

**Abstract**—A smart grid state estimation problem is studied, in which the control center (CC) sends state measurement requests to multiple sensors. Each sensor is capable of measuring multiple system parameters, and the CC asks for a particular system parameter from each sensor at each time slot. Delay constrained linear transmission of state measurements over orthogonal fading channels is studied. The total mean-square error (MSE) distortion is studied for Gaussian system parameters under an average power constraint. For a particular symmetric scenario, a round-robin scheduling algorithm and optimal linear transmission (LT) schemes are presented. It is shown that, as opposed to the case of a single parameter measurement, the performance of the proposed LT strategy improves as the delay constraint is relaxed, and achieves the Shannon lower bound under certain matching conditions between the source and the channel statistics.

## I. INTRODUCTION

Closely monitoring the state of a smart grid (SG) is essential to manage and control the network efficiently, and to detect and reduce the impact of failures [1]. Sensors are deployed throughout the grid to track sensitive system parameters, such as voltage, current, active/reactive power values, as well as temperature and other physical parameters [2]. Wireless sensors are attractive for grid monitoring due to their flexibility, low cost and easy maintenance [3]. On the other hand, near real-time monitoring of the grid imposes strict delay constraints on the transmission of sensor measurements to the control center (CC); and hence, it is not practical to consider advanced compression and channel coding techniques. Therefore, we focus on linear transmission (LT) of sensor measurements, which not only satisfies the delay constraints, but also limits the complexity of the encoding scheme; and hence, the cost and energy requirements of the sensors.

We consider  $N$  wireless sensors, each of which is capable of measuring  $J$  distinct parameters locally. Time is discretized into time slots (TSs), and at each TS, measurement of a particular parameter is requested from each sensor by the CC. The sensors learn these requests causally, but their statistics, i.e., the probability with which each measurement is requested, is known in advance. There are  $N$  orthogonal channels available for the transmission of the requested measurements to the CC. These channels are scheduled to sensors in advance; which cannot depend on the realizations of the measurements or the channel states. We focus on a symmetric model, in which the

statistics of the local measurements, the request probabilities for each parameter, as well as the fading channel distributions are the same for all the sensors. We study the minimum mean square error (MSE) distortion achievable by LT under delay constraints.

LT of Gaussian sources over Gaussian channels has been previously studied in the literature. In [4], the optimal linear vector encoding scheme that matches an  $r$ -dimensional Gaussian signal to a  $k$ -dimensional additive white Gaussian noise (AWGN) vector channel under an average power constraint is characterized. In [5], LT of a Gaussian source over a fading AWGN channel is studied. It is shown that the optimal linear encoding performance can be achieved by symbol-by-symbol processing, and increasing the block length does not provide any improvement, as opposed to nonlinear coding schemes. In [6], linear encoding-decoding strategies for the transmission of the noisy vector state measurements over a fading AWGN channel is studied under diagonal and general observation matrices. LT of vector Gaussian sources over multi-antenna static and fading channels is studied in [7].

In our model  $J$  system measurements are not all available at the sensors at each TS, and instead, the source to be transmitted can be considered as a Gaussian mixture. We propose a round-robin scheduling algorithm depending on the delay constraint, find the corresponding optimal LT scheme, and characterize the achievable average sum distortion subject to delay and power constraints. In contrast to [5], we observe that in our scenario the achievable average sum distortion diminishes as the delay constraint is relaxed. We then consider the extreme case when the delay constraint is removed, and show that LT achieves the Shannon lower bound under certain matching conditions between the source and channel statistics.

## II. SYSTEM MODEL

Consider  $N$  sensors, each of which is capable of measuring  $J$  distinct state parameters locally. The  $j$ th state parameter is denoted by  $S_j$ , and is modeled as a zero-mean Gaussian random variable (r.v.) with variance  $\sigma_j^2$ , i.e.,  $S_j \sim \mathcal{N}(0, \sigma_j^2)$ , for  $j \in [1, J]$ , where  $[1, J]$  denotes the set  $\{1, 2, 3, \dots, J\}$ . Each state parameter is independent and identically distributed (i.i.d.) over time and over sensors. At each TS, which corresponds to one use of the network, each sensor measures one state parameter requested by the CC. Let r.v.  $M \in [1, J]$  denote the requested state parameter.  $M$  is also i.i.d. over time

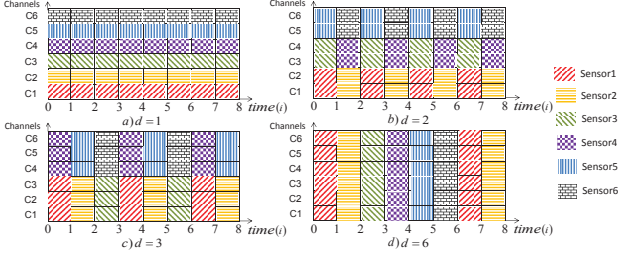


Fig. 1. Illustration of the proposed channel scheduling scheme for different delay constraints.

and sensors, with probability distribution  $\Pr\{M = m\} = p_M(m)$ .

The maximum delay in transmitting a measurement to the CC is  $d \in \mathbb{N}^+$ , which is same for all the sensors and parameters. There are  $N$  orthogonal fading channels available in the system. Let  $\mathbf{h} = [h_1, \dots, h_N]$  denote the channel vector at any TS. The entries of  $\mathbf{h} \in \mathbb{R}^N$  are i.i.d. with probability distribution  $p_H(h)$ . We assume a fast-fading model, i.e.,  $\mathbf{h}$  changes at each TS in an i.i.d. fashion. The channels have AWGN with zero-mean and unit-variance. We note that we consider a symmetric system model among sensors, namely, the statistics of the system parameters  $S_j$ , measurement requests  $M$ , and channel states  $h$  are all i.i.d. over sensors.

Scheduling of channels to sensors is done in advance; i.e., it cannot depend on the realizations of the measurements or the channel states. At each scheduled TS for transmission, each sensor transmits all its samples that have been taken within the last  $d$  TSs over the channels assigned to it. We consider an average power constraint of  $P$  at each sensor. The goal is to have an estimate of each requested measurement at the CC before the delay constraint expires. The performance measure is the total MSE distortion for the requested measurements.

#### A. Scheduling Algorithm

We consider a round-robin scheduling algorithm. Given a delay constraint  $d$ , assuming  $N$  is an integer multiple of  $d$ , we group sensors into  $N/d$  groups. Each group is assigned  $d$  orthogonal channels. Each sensor transmits once every  $d$  TSs, using all the  $d$  channels assigned to its own group, and transmits all its measurements from the last  $d$  TSs. This round-robin scheduling of channels provides additional degrees-of-freedom to the sensors to match their measurements to a larger number of channels at each transmission round. Fig. 1 depicts an example of how  $N = 6$  channels are scheduled to  $N = 6$  sensors for different delay constraints  $d = \{1, 2, 3, 6\}$ . Notice that for  $d = 1$ , one channel is assigned to each sensor for all the TSs; hence, the sensor has no control on matching the measurements to the channel states. On the other hand, for  $d > 1$ , at each transmission round, a sensor can reorder its measurements to match them to the available channels in an optimal manner, or transmit their linear combinations.

We analyze the system performance for a single sensor. Due to symmetry, results will apply to all the sensors. For a delay constraint  $d$ , the sensor collects  $d$  measurements to

be transmitted over  $d$  orthogonal fading AWGN channels. Let  $\mathbf{m} = \langle m_1, \dots, m_d \rangle$  be the sequence of  $d$  most recent requested measurements including the current TS. The encoder input vector is thus  $\mathbf{u} = [s_{m_1}, s_{m_2}, \dots, s_{m_d}]$ , where each entry is the measured value of the requested parameter, i.e.,  $s_{m_2}$  is the measured value of the  $m_2$ th parameter.

The signal received at the CC is given by  $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z}$ , where  $\mathbf{x}$  and  $\mathbf{y}$  are the length- $d$  channel input and output vectors, respectively, and  $\mathbf{z}$  is the length- $d$  i.i.d. AWGN noise vector with  $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ .  $\mathbf{H}$  is a  $d \times d$  diagonal channel matrix with i.i.d. fading states. The encoder maps its measurement vector  $\mathbf{u}$ , to a channel input vector  $\mathbf{x}$ , using an encoding function  $f: \mathbb{R}^d \rightarrow \mathbb{R}^d$ , that satisfies the average power constraint  $P$ :

$$\frac{1}{d} \mathbb{E}_{M, H} [\text{Tr} \{ \mathbb{E} [\mathbf{x} \mathbf{x}^T] \}] \leq P.$$

The receiver estimates the measurement vector as  $\hat{\mathbf{u}} = g(\mathbf{y})$ ,  $g: \mathbb{R}^d \rightarrow \mathbb{R}^d$ . The average sum quadratic distortion measure is given by:

$$\bar{D} = \frac{1}{d} \mathbb{E}_{M, H} [\text{Tr} \{ \mathbb{E}_{S, Z} [(\mathbf{u} - \hat{\mathbf{u}})(\mathbf{u} - \hat{\mathbf{u}})^T] \}],$$

We assume that both the encoder and the decoder know the instantaneous channel state, and the statistics of the measured parameters,  $\sigma_m^2$ , the measurement requests,  $p_M(m)$ , and the channels,  $p_H(h)$ .

We are interested in the minimum MSE distortion for the state measurements  $\bar{D}^*$ , when  $f$  and  $g$  are restricted to be linear. Hereafter,  $f$  and  $g$  are represented by  $d \times d$  matrices  $\mathbf{F}$  and  $\mathbf{G}$ , respectively.

### III. STRICT DELAY CONSTRAINT

We first consider a strict delay constraint of  $d = 1$ . In this case, the encoding function is a scalar  $f$ . Given the encoding function, the decoding function  $g$ , that minimizes the MSE for Gaussian r.v.s is the linear MMSE estimator [8], and thus also a scalar. Then, the MSE  $\bar{D}$ , and the average power  $\bar{P}$ , as a function of the fading state  $h$ , and the variance of the measurement  $\sigma_m^2$ , are given by:

$$\bar{D} = \sum_{m=1}^J p_M(m) \int \frac{\sigma_m^2}{h^2 f^2 \sigma_m^2 + 1} p_H(h) dh, \quad (1)$$

$$\bar{P} = \sum_{m=1}^J p_M(m) \int f^2 \sigma_m^2 p_H(h) dh. \quad (2)$$

We define  $w \triangleq \frac{\sigma_m}{h}$  and  $l \triangleq \frac{1}{h \sigma_m}$ . The optimal linear encoding function  $f^*$  is found as the solution to the convex optimization problem  $\bar{D}^* \triangleq \min \bar{D}$ , subject to the average power constraint  $\bar{P} \leq P$ , as follows:

$$f^* = l \sqrt{\left[ \frac{\lambda^*}{l} - 1 \right]^+}, \quad (3)$$

where  $\lambda^*$  is the optimal Lagrange multiplier, such that  $\bar{P} \leq P$ .

Then the optimal power allocation and the corresponding distortion are given by:

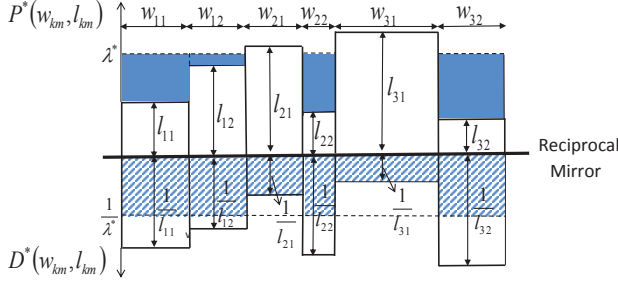


Fig. 2. Water-filling reflected on a reciprocal mirror.

$$P^*(w, l) = w [\lambda^* - l]^+, \quad (4)$$

$$D^*(w, l) = w \min \left( \frac{1}{\lambda^*}, \frac{1}{l} \right). \quad (5)$$

In Fig. 2, we present a graphical interpretation of the optimal power allocation and the corresponding distortion for  $J = 2$  parameters with variances  $\{\sigma_1^2, \sigma_2^2\}$ , which are requested with probabilities  $\{p_M(1), p_M(2)\}$ , respectively, and a discrete fading channel with three states, where the  $k$ th state  $\hat{h}_k$ , is observed with probability  $p_H(\hat{h}_k)$ ,  $k = 1, 2, 3$ . Fig. 2 depicts rectangles that are placed upon a mirror surface and their reciprocally scaled images below. Rectangles represent all possible source-channel pairs  $\{\sigma_m, \hat{h}_k\}$ , where  $l_{km} = \frac{1}{\hat{h}_k \sigma_m}$  and  $w_{km} = \frac{\sigma_m}{\hat{h}_k}$  indicate the length and width of the rectangles, respectively. While the optimal power allocation is obtained by water-filling across the rectangles placed upon the mirror, corresponding distortion values are found by simply looking at the reflection of this on the mirror, which takes the reciprocal of the heights. We call this *water-filling reflected on a reciprocal mirror*.

#### IV. RELAXED DELAY CONSTRAINT

Next, we consider a delay constraint  $1 < d \leq N$ . We characterize the corresponding optimal LT, and the achievable minimum MSE distortion. Given the linear encoding matrix  $\mathbf{F}$ , diagonal channel matrix  $\mathbf{H}$ , and the covariance matrix of the measurement vector  $\mathbf{C}_u$ , the optimal decoding function is again the linear MMSE estimator matrix  $\mathbf{G}$ . Then, the MSE distortion  $\bar{D}$ , and the average power  $\bar{P}$  are given by:

$$\bar{D} = \frac{1}{d} \text{Tr} \{ \mathbb{E}_{\mathbf{M}, \mathbf{H}} [\mathbf{C}_u - \mathbf{C}_u \mathbf{F}^T \mathbf{H}^T \Phi^{-1} \mathbf{H} \mathbf{F} \mathbf{C}_u] \}, \quad (6)$$

$$\bar{P} = \frac{1}{d} \text{Tr} \{ \mathbb{E}_{\mathbf{M}, \mathbf{H}} [\mathbf{F} \mathbf{C}_u \mathbf{F}^T] \}, \quad (7)$$

where  $\mathbf{C}_u = \mathbb{E}[\mathbf{u} \mathbf{u}^T]$  and  $\Phi \triangleq \mathbf{H} \mathbf{F} \mathbf{C}_u \mathbf{F}^T \mathbf{H}^T + \mathbf{I}$ .

For a set of static parallel AWGN channels and Gaussian vector sources, the linear encoding matrix  $\mathbf{F}^*$  that minimizes the MSE distortion  $\bar{D}$ , subject to an average power constraint  $P$  is derived in [4]. It is shown that the optimal linear encoding matrix transmits one measurement over each channel. The optimal mapping between channels and measurements is as follows: We first reorder the measurements  $\mathbf{u}$ , to get

$\bar{\mathbf{u}} = [s_{m(1)}, s_{m(2)}, \dots, s_{m(d)}]$ , where the variances of the ordered measurements satisfy  $\sigma_{m(1)}^2 \leq \sigma_{m(2)}^2 \leq \dots \leq \sigma_{m(d)}^2$ , and reorder the fading states  $\mathbf{h} = [h_1, h_2, \dots, h_d]$ , such that the reordered channel states  $\bar{\mathbf{h}} = [\hat{h}_{(1)}, \hat{h}_{(2)}, \dots, \hat{h}_{(d)}]$  satisfy  $h_{(1)} \leq h_{(2)} \leq \dots \leq h_{(d)}$ . Then, the optimal linear encoding matrix  $\bar{\mathbf{F}}^*$  is diagonal  $\bar{\mathbf{F}}^* = \text{diag} [f_{(1)}, f_{(2)}, \dots, f_{(d)}]$ , and its diagonal entries can be found solving the following convex optimization problem:

$$\begin{aligned} \bar{D}^* \triangleq \min_{f_{(1)}, \dots, f_{(d)}} \mathbb{E}_{\mathbf{M}_{(t)}, \mathbf{H}_{(t)}} & \left[ \frac{1}{d} \sum_{t=1}^d \frac{\sigma_{m(t)}^2}{h_{(t)}^2 f_{(t)}^2 \sigma_{m(t)}^2 + 1} \right] \\ \text{s.t. } \mathbb{E}_{\mathbf{M}_{(t)}, \mathbf{H}_{(t)}} & \left[ \frac{1}{d} \sum_{t=1}^d f_{(t)}^2 \sigma_{m(t)}^2 \right] \leq P, \end{aligned} \quad (8)$$

where the expectations are taken over  $M_{(t)}$  and  $H_{(t)}$ . The order statistics of the  $t$ -th smallest element of  $\{m_1, m_2, \dots, m_d\}$ , which are i.i.d. random samples of  $M$ , are denoted by the r.v.  $M_{(t)} \in [1, J]$  with distribution  $p_{M_{(t)}}(m)$ . Without loss of generality, we assume that ordering i.i.d. random samples of  $M_{(t)}$  in ascending order, i.e.,  $m_{(1)} \leq m_{(2)} \leq \dots \leq m_{(d)}$ , implies ordering the measurements in ascending variances, i.e.,  $\sigma_{m(1)}^2 \leq \sigma_{m(2)}^2 \leq \dots \leq \sigma_{m(d)}^2$ . Similarly, the order statistics of the  $t$ -th smallest element of  $\{h_1, h_2, \dots, h_d\}$ , the i.i.d. random fading states of the channels, is denoted by the r.v.  $H_{(t)} \in \mathbb{R}$  with distribution  $p_{H_{(t)}}(h)$ .

Since the objective function and the inequality constraints are convex functions of  $f_{(t)}^2$ , from the KKT conditions we obtain the optimal linear encoding matrix as:

$$f_{(t)}^* = l_{(t)} \sqrt{\left[ \frac{\delta^*}{l_{(t)}} - 1 \right]^+}, \quad (9)$$

where  $l_{(t)} = \frac{1}{h_{(t)} \sigma_{m(t)}}$ , and  $\delta^*$  is the optimal Lagrange multiplier, such that  $\bar{P} \leq P$  under the delay constraint  $d$ .

Similar to strict delay constraint, the optimal power allocation and the corresponding distortion can be found by water-filling, and its reflection on a reciprocal mirror. The optimal  $\delta^*$  depends on  $p_{M_{(t)}}(m)$  and  $p_{H_{(t)}}(h)$ , which can be found explicitly using the order statistics.

#### V. NO DELAY CONSTRAINT

Another extreme scenario for our system model is obtained when the delay constraint is completely removed. We first state the theoretical performance bound without any delay or complexity constraints, and prove that this lower bound can sometimes be achieved by LT.

##### A. The Shannon Lower Bound

Shannon's well known source-channel separation theorem states that the optimal end to end distortion is achieved by concatenating the optimal source and channel codes, when there is no delay or complexity constraints, and if the source and channel distributions are ergodic. For the symmetric sensor model studied in this paper, the Shannon lower bound is

achieved by allocating one channel to each sensor. The sensor can transmit to the CC at the ergodic capacity,  $\bar{C}_e$ , of the underlying fading channel, while the achieved distortion,  $\bar{D}_e$ , is found by evaluating the distortion-rate function for our source model at the ergodic channel capacity.

Since the transmitter and receiver both know the channel state, the ergodic capacity in terms of the optimal power allocation scheme  $P_e^*(h)$  is given by:

$$\bar{C}_e \triangleq \mathbb{E}_H \left[ \frac{1}{2} \log (1 + h^2 P_e^*(h)) \right], \quad (10)$$

where  $P_e^*(h)$  is found by the water-filling algorithm as  $P_e^*(h) = [\alpha^* - 1/h^2]^+$ , such that  $\alpha^*$  satisfies  $\mathbb{E}_H [P_e^*(h)] \leq P$ .

The distortion-rate function for a Gaussian source  $S_m \sim N(0, \sigma_m^2)$ , each of which is observed with probability  $p_M(m)$  for  $m \in [1, J]$ , is defined in terms of the optimal rate allocation function  $R_e^*(\sigma_m)$  as:

$$\bar{D}_e \triangleq \mathbb{E}_M \left[ \sigma_m^2 2^{-2R_e^*(\sigma_m)} \right]. \quad (11)$$

$R_e^*(\sigma_m)$  and the distortion  $D_e^*(\sigma_m)$  are:

$$R_e^*(\sigma_m) = \frac{1}{2} \left[ \log \left( \frac{\sigma_m^2}{\beta^*} \right) \right]^+, \quad (12)$$

$$D_e^*(\sigma_m) = \min (\beta^*, \sigma_m^2), \quad (13)$$

where  $\beta^*$  is chosen such that  $\mathbb{E}_M [R_e^*(\sigma_m)] \leq \bar{C}_e$ .

The Shannon lower bound is not tight when the encoder is limited to LT, even if the delay constraint is removed. However, we show in the following theorem, which is stated here without a proof due to space limitation, that, the Shannon lower bound is tight when the channels follow a discrete probability distribution with the same number of states as the number of different parameters,  $J$ , and the channel states and the source variances satisfy a certain matching condition.

Assume that the channel state distribution is discrete with  $J$  states, and without loss of generality, channels states and the source variances are ordered as,  $\hat{h}_1 > \hat{h}_2 > \dots > \hat{h}_J$ , and  $\sigma_1^2 > \sigma_2^2 > \dots > \sigma_J^2$ .

**Theorem 1.** *For the discrete fading channel model introduced above, under no delay constraint, i.e.,  $d \rightarrow \infty$ , if the source variances and the channel states satisfy  $\frac{\sigma_1^2}{\hat{h}_1} = \dots = \frac{\sigma_J^2}{\hat{h}_J}$ , and  $p_M(m) = p_H(\hat{h}_m)$ , for all  $m \in [1, J]$ , then the Shannon lower bound can be achieved by LT.*

## VI. NUMERICAL RESULTS

We illustrate the performance of LT for strict, relaxed and no delay constraint scenarios through numerical results. Fig. 3(a) shows the achievable average sum distortion versus the average power for  $J = 4$  Gaussian system parameters with variances  $\{625, 225, 64, 4\}$ , each of which is requested with equal probability, i.e.,  $p_M(m) = 0.25$ , for  $m \in [1, 4]$ . The discrete fading channel has four states  $\{25, 15, 8, 2\}$ , each of which is observed with equal probability, i.e.,  $p_H(\hat{h}_k) = 0.25$ , for  $k \in [1, 4]$ . We observe that the achievable average sum

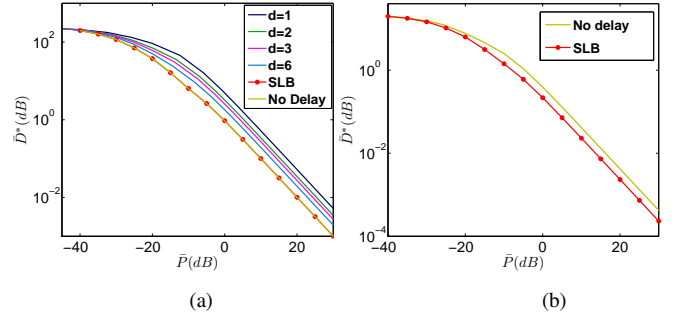


Fig. 3. Achievable average sum distortion with respect to average power.

distortion decreases as the delay constraint is relaxed. In addition, Fig. 3(a) illustrates that LT under no delay constraint achieves the Shannon lower bound, as claimed by Theorem 1.

Fig. 3(b) shows the MSE distortion versus the average power for  $J = 4$  parameters with variances  $\{30, 25, 15, 2\}$ , which are requested with probabilities  $\{0.1, 0.6, 0.2, 0.1\}$ , respectively. Considering the same discrete fading channel as above, we observe that the LT in this case does not meet the Shannon lower bound even when the delay constraint is relaxed; however, the gap between the two is quite small.

## VII. CONCLUSIONS

We have studied the delay constrained linear encoding of multiple state measurements from  $N$  sensors to a CC over orthogonal fading AWGN channels. We have employed a channel scheduling algorithm that schedules the available channels to sensors in a round-robin fashion in order to provide the maximum degrees-of-freedom to each sensor when matching its measurements to the orthogonal fading channels. We have characterized the achievable average sum distortion for varying delay constraints, and showed that the average sum distortion decreases as the delay constraint is relaxed. We have also characterized the Shannon lower bound when the delay constraint is removed, and shown that LT achieves this lower bound under certain matching conditions between the source and the channel statistics.

## REFERENCES

- [1] A. Ipakchi and F. Albuyeh, "Grid of the future," *IEEE Power Energy Mag.*, vol. 7, no. 2, pp. 52–62, Mar.-Apr. 2009.
- [2] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. New York, NY: CRC Press, 2000.
- [3] V. C. Gungor, B. Lu, and G. P. Hancke, "Opportunities and challenges of wireless sensor networks in smart grid," *IEEE Trans. Industrial Electronics*, vol. 57, pp. 3557–3564, Oct. 2010.
- [4] K. H. Lee and D. P. Petersen, "Optimal linear coding for vector channels," *IEEE Trans. Commun.*, vol. 24, no. 12, pp. 1283–1290, Dec. 1976.
- [5] J. J. Xiao, Z. Q. Luo, and N. Jindal, "Linear joint source-channel coding for gaussian sources through fading channels," in *Proc. IEEE Global Telecommun. Conf.*, San Francisco, CA, Dec. 2006.
- [6] J. Geng, H. Li, and L. Lai, "Smart grid system state measurement estimation over wireless channels," in *Proc. 46th Inform. Sci. Sys. Conf.*, Princeton, NJ, Mar. 2012.
- [7] I. E. Aguerri and D. Gunduz, "Linear transmission of correlated Gaussian sources over MIMO channels," in *Proc. IEEE Int. Symp. on Wireless Commun. Systems (ISWCS)*, Ilmenau, Germany, Aug. 2013.
- [8] A. E. Gamal and Y. H. Kim, *Network Information Theory*. Cambridge University Press, 2011.